From Newton to Runge Kutta to ODEs!

PHYS 250 (Autumn 2018) – Lecture 10

David Miller

Department of Physics and the Enrico Fermi Institute
University of Chicago

November 6, 2018

Outline

Reminders from last time

Looked at two primary and exemplary methods for root finding, which is part of the foundation of optimization and differential equation solving.

Fundamental root finding methods

- **Bisection method (aka "incremental search"):**
  - **PROs:** exceptionally simple and requires no knowledge of the function whose roots are sought
  - **CONs:** doesn't use the potentially very useful knowledge of the roots that are sought
- **Newton's Method:**
  - **PROs:** converges much faster than bisection
  - **CONs:** requires a calculation or estimation of the first derivative of the function

Today, we will expand on these algorithms and go several steps further.

Outline

Recall the description of Newton's method

Recall that Newton's method uses the Taylor expansion

$$F(x_0) = F(x+\delta) \approx F(x) + \delta F'(x) + \frac{1}{2}\delta^2 F''(x) + \mathcal{O}(\delta^3) \tag{1}$$

to inform the use of a linear approximation $\delta \approx \Delta$ where

$$\Delta = -\frac{F(x)}{F'(x)} \tag{2}$$

That gives way to an iterative approach that updates the estimate of the position of the root of $F(x)$ as being at x_{i+1}:

$$x_{i+1} = x_i - \frac{F(x_i)}{F'(x_i)} \tag{3}$$

The iteration stops after j iterations when

$$|x - x_j| \leq \epsilon \tag{4}$$

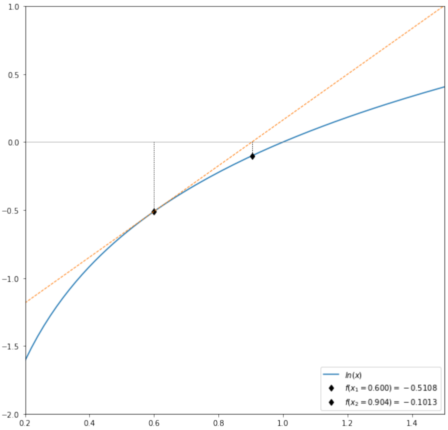
Precision of Newton's method

For any estimate x_i of the method, the error, E_i is the difference between the true root x and the estimate:

$$E_i = x - x_i \tag{5}$$

Merely by inspecting the design of Newton's method, you can see that the precision of the estimate for a subsequent iteration will be given by

$$\begin{aligned} E_{i+1} &= E_i + \frac{F(x)}{F'(x)} && (6) \\ &= -\frac{F''(x)}{2F'(x)} E_i^2 && (7) \end{aligned}$$

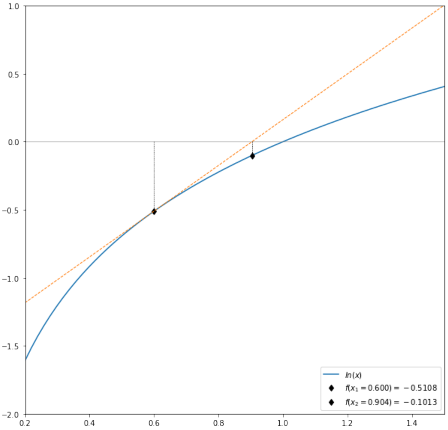


Convergence of Newton's method

Consequently, Newton's method **converges quadratically**

- the error is the square of the error in the previous step)
- the number of significant figures is roughly doubled in every iteration, provided that x_i is **sufficiently** close to the root.

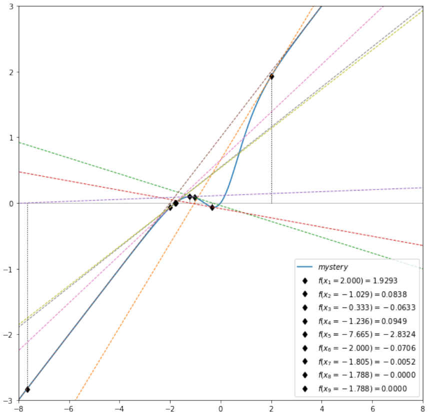
However, a critical assumption is that $F'(x) \neq 0$; for all $x \in I$, where I is the interval $[x - r, x + r]$ for some $r \geq |x - x_0|$ and x is the true root and x_0 was the starting point.



Pathologies and divergent scenarios

That is definitely not always the case. Let's look at a pathological example. Here is a fun mystery function that I cooked up (since you need to do something similar on your homework):

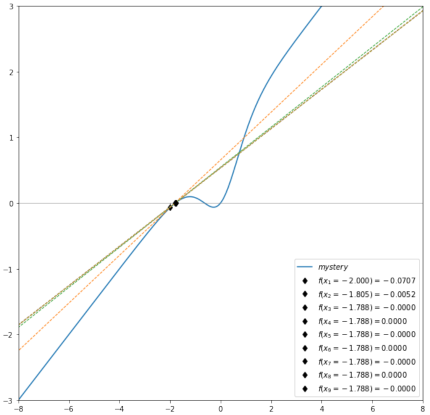
- $x_0 = 2.000$, **7 iterations**
- $x_0 = -2.000$, **2 iterations**
- $x_0 = 3.000$, **2 iterations**
- $x_0 = -1.214$, **4 iterations**
- Slight modification: $x_0 = -1.213$, **no convergence**
- Slight modification: $x_0 = 3.000$, **no convergence**



Pathologies and divergent scenarios

That is definitely not always the case. Let's look at a pathological example. Here is a fun mystery function that I cooked up (since you need to do something similar on your homework):

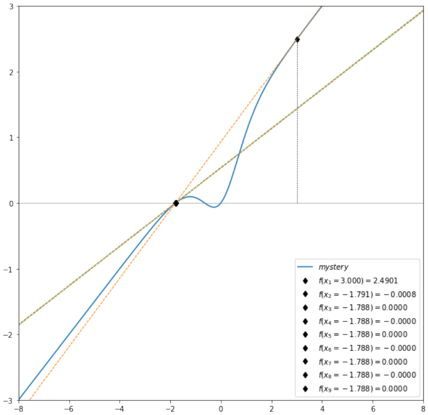
- $x_0 = 2.000$, **7 iterations**
- $x_0 = -2.000$, **2 iterations**
- $x_0 = 3.000$, **2 iterations**
- $x_0 = -1.214$, **4 iterations**
- Slight modification: $x_0 = -1.213$, **no convergence**
- Slight modification: $x_0 = 3.000$, **no convergence**



Pathologies and divergent scenarios

That is definitely not always the case. Let's look at a pathological example. Here is a fun mystery function that I cooked up (since you need to do something similar on your homework):

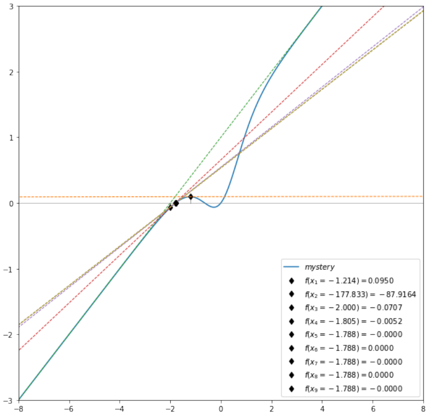
- $x_0 = 2.000$, **7 iterations**
- $x_0 = -2.000$, **2 iterations**
- $x_0 = 3.000$, **2 iterations**
- $x_0 = -1.214$, **4 iterations**
- Slight modification: $x_0 = -1.213$, **no convergence**
- Slight modification: $x_0 = 3.000$, **no convergence**



Pathologies and divergent scenarios

That is definitely not always the case. Let's look at a pathological example. Here is a fun mystery function that I cooked up (since you need to do something similar on your homework):

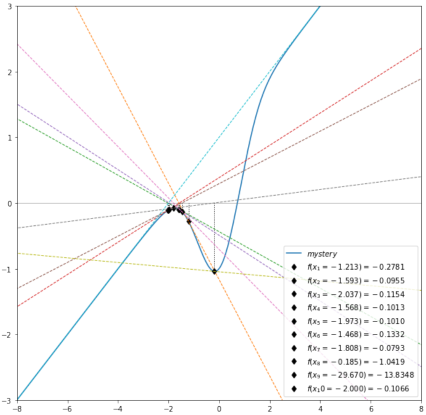
- $x_0 = 2.000$, **7 iterations**
- $x_0 = -2.000$, **2 iterations**
- $x_0 = 3.000$, **2 iterations**
- $x_0 = -1.214$, **4 iterations**
- Slight modification: $x_0 = -1.213$, **no convergence**
- Slight modification: $x_0 = 3.000$, **no convergence**



Pathologies and divergent scenarios

That is definitely not always the case. Let's look at a pathological example. Here is a fun mystery function that I cooked up (since you need to do something similar on your homework):

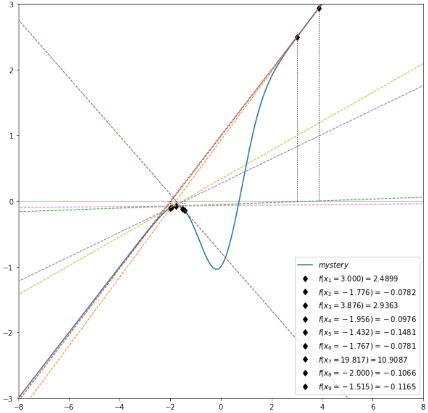
- $x_0 = 2.000$, **7 iterations**
- $x_0 = -2.000$, **2 iterations**
- $x_0 = 3.000$, **2 iterations**
- $x_0 = -1.214$, **4 iterations**
- Slight modification:
  $x_0 = -1.213$, **no convergence**
- Slight modification: $x_0 = 3.000$, **no convergence**



Pathologies and divergent scenarios

That is definitely not always the case. Let's look at a pathological example. Here is a fun mystery function that I cooked up (since you need to do something similar on your homework):

- $x_0 = 2.000$, **7 iterations**
- $x_0 = -2.000$, **2 iterations**
- $x_0 = 3.000$, **2 iterations**
- $x_0 = -1.214$, **4 iterations**
- Slight modification: $x_0 = -1.213$, **no convergence**
- Slight modification: $x_0 = 3.000$, **no convergence**



Backtracking

In the last examples above we have a case where the search falls into the pathology of a situation where the initial guess was not **sufficiently close** to the root. an "infinite" loop without ever getting there.

A solution to this problem is called **backtracking**.

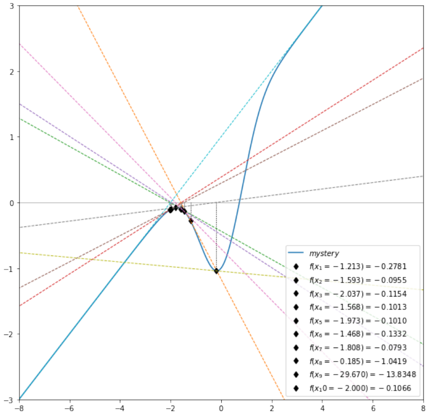
Backtracking

In cases where the new guess $x_0 + \Delta x$ leads to an increase in the magnitude of the function, $|f(x_0 + \Delta x)|^2 > |f(x_0)|^2$, you should backtrack somewhat and try a smaller guess, say, $x_0 + \Delta x/2$. If the magnitude of f still increases, then you just need to backtrack some more, say, by trying $x_0 + \Delta x/4$ as your next guess, and so forth.

Pathological case fixed with backtracking

Fixing the pathological example with backtracking:

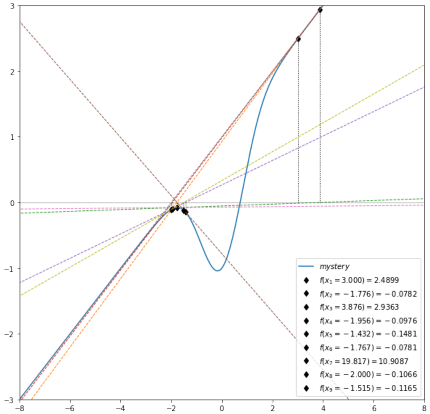
- $x_0 = -1.213$, **no convergence**
- $x_0 = 3.000$, **no convergence**
- $x_0 = 1.500$, **3 iterations**



Pathological case fixed with backtracking

Fixing the pathological example with backtracking:

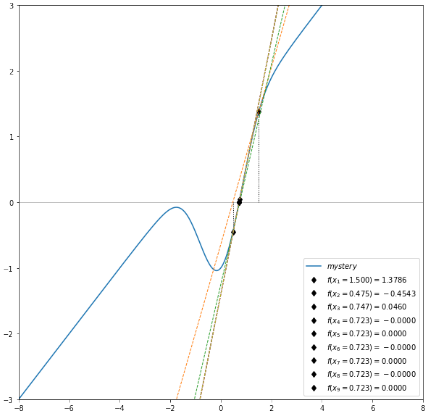
- $x_0 = -1.213$, **no convergence**
- $x_0 = 3.000$, **no convergence**
- $x_0 = 1.500$, **3 iterations**



Pathological case fixed with backtracking

Fixing the pathological example with backtracking:

- $x_0 = -1.213$, **no convergence**
- $x_0 = 3.000$, **no convergence**
- $x_0 = 1.500$, **3 iterations**



Multidimensional problems

Up to this point, we have confined our attention to solving the single equation $F(x) = 0$. Let us now consider the n-dimensional version of the same problem, namely

$$\vec{F}(\vec{x}) = 0 \tag{8}$$

where we allow for a vector of functions $\vec{F} = \{f_1(\vec{x}), f_2(\vec{x}), ..., f_n(\vec{x})\}$, and $\vec{x} = \{x_1, x_2, ..., x_n\}$.

The solution of n simultaneous, nonlinear equations is a much more formidable task than finding the root of a single equation. The trouble is the lack of a reliable method for bracketing the solution vector $\vec{x}$. Therefore, we cannot always provide the solution algorithm with a good starting value of x, unless such a value is suggested by the physics of the problem.

Newton's method is the workhorse here!

Reminder of the general problem

Start by considering each one of the n functions, $f_n(x)$ separately:

$$
\begin{aligned}
f_i(\vec{x}) &= f_i(\vec{a}) + \sum_j^n \frac{\partial f_i}{\partial x_j}|_{\vec{x}} \Delta x_j + \frac{1}{2!} \sum_j^n \sum_k^n \frac{\partial^2 f_i}{\partial_j \partial_k}|_{\vec{x}} \Delta x_j \Delta x_k && (9) \\
&= f_i(\vec{a}) + (\vec{x} - \vec{a})^{\mathrm{T}} \nabla f(\vec{a}) + \frac{1}{2!} (\vec{x} - \vec{a})^{\mathrm{T}} \mathbf{H}(\vec{a}) (\vec{x} - \vec{a}) && (10)
\end{aligned}
$$

where **H** is the **Hessian matrix**, describing the **curvature** of $f(\vec{x})$ by

Jacobian matrix

Something

- Something
  - Something else

Lastly

Examples

Something

- Something
  - Something else

Lastly

Comments on matrix computing and manipulations

Something

- Something
  - Something else

Lastly

Finite difference approximation

Something

- Something
  - Something else

Lastly

Derivatives by Interpolation

Something

- Something
    - Something else

Lastly

Outline

Initial vs. boundary values

Something

- Something
  - Something else

Lastly

Recall the Taylor series approach generally

Something

- Something
  - Something else

Lastly

Avoiding repeated differentiation: Runge?Kutta

Something

- Something
  - Something else

Lastly

Precision

Something

- Something
  - Something else

Lastly

Higher order calculations: `rk4`

Something

- Something
  - Something else

Lastly

Outline

The issue

In an initial value problem we were able to start at the point where the initial values were given and march the solution forward as far as needed. This technique does not work for boundary value problems, because there are not enough starting conditions available at either endpoint to produce a unique solution.

- Something
  - Something else

Lastly

Recall the Taylor series approach generally

Something

- Something
  - Something else

Lastly